



OPEN

## Estimating generation time of SARS-CoV-2 variants in Italy from the daily incidence rate

Eugenio Lippiello<sup>1✉</sup>, Giuseppe Petrillo<sup>2</sup>, Silvio Baccari<sup>1</sup> & Lucilla de Arcangelis<sup>1</sup>

The identification of the transmission parameters of a virus is fundamental to identify the optimal public health strategy. These parameters can present significant changes over time caused by genetic mutations or viral recombination, making their continuous monitoring fundamental. Here we present a method, suitable for this task, which uses as unique information the daily number of reported cases. The method is based on a time since infection model where transmission parameters are obtained by means of an efficient maximization procedure of the likelihood. Applying the method to SARS-CoV-2 data in Italy, we find an average generation time  $\bar{z} = 3.2 \pm 0.8$  days, during the temporal window when the majority of infections can be attributed to the Omicron variants. At the same time we find a significantly larger value  $\bar{z} = 6.2 \pm 1.1$  days, in the temporal window when spreading was dominated by the Delta variant. We are also able to show that the presence of the Omicron variant, characterized by a shorter  $\bar{z}$ , was already detectable in the first weeks of December 2021, in full agreement with results provided by sequences of SARS-CoV-2 genomes reported in national databases. Our results therefore show that the novel approach can indicate the existence of virus variants, resulting particularly useful in situations when information about genomic sequencing is not yet available. At the same time, we find that the standard deviation of the generation time does not significantly change among variants.

SARS-CoV-2, as other viruses, are continuously evolving because of genetic mutations or viral recombination. These changes can strongly affect transmission parameters<sup>1</sup> inducing important differences in the virus spreading. In particular a reduction of the generation time  $z$ , i.e. the time difference between the dates of infection of successive cases in a transmission chain, leads to an increased epidemic growth rate, even for unaltered reproduction number  $R_0$ . Furthermore, an accurate estimate of the mean value of the generation time  $\bar{z}$  is fundamental to establish the optimal duration of the quarantine period.

Elegant methods based on log-likelihood maximization have been recently developed<sup>2–5</sup> to obtain the average value  $\bar{z}$  of the generation time. However, very often  $\bar{z}$  is identified with  $\bar{s}$ , defined as the mean value of the serial interval<sup>4</sup>, which is the difference in timing of symptom onset in a pair of a primary and its secondary case. The measurement of  $\bar{s}$ , indeed, differently from the measurement of  $\bar{z}$ , can be directly obtained from the reconstruction of the contact network. This information, combined with the results provided by genomic sequencing, provides an estimate of the mean value of serial intervals of each specific variant<sup>6</sup>. Nevertheless, it is important to remark<sup>7–10</sup> that the value of  $\bar{s}$ , obtained from contact tracing can be significantly different than the “intrinsic” value of  $\bar{z}$ . This occurs, in particular, when the structure of the contact network fastly changes in time, as for instance in presence of non-pharmaceutical interventions. Here “intrinsic” refers to the quantity measured in the ideal case of a fully susceptible, homogeneously mixed population<sup>7</sup>, and therefore independent of the specific conditions of the epidemiological setting from which it is inferred.

In this study we will show that the intrinsic value of  $\bar{z}$  can be obtained by means of a completely data driven procedure. The main observation is that, if the value of  $\bar{z}$  affects the future evolution of the number of infected cases, its value could be potentially extracted from the previous evolution of the virus spreading. More precisely we use the method recently developed by us<sup>5</sup> to extract  $\bar{z}$  directly from the daily series of incidence rate  $I(t)$ , i.e. the number of infected individuals at the calendar time  $t$ . The method is based on the non-trivial dependence on  $\bar{z}$  of the Log-Likelihood function  $LL$ , which measures the overlap between the measured  $I(t)$  and the expected one, according to a time since infection model<sup>11,12</sup>. We show that when two variants with differences in  $\bar{z}$  are simultaneously present in the sample,  $LL(\bar{z})$  presents two distinct peaks in correspondence to the mean value of

<sup>1</sup>Department of Mathematics and Physics, University of Campania “L. Vanvitelli”, Viale Lincoln, 5, 81100 Caserta, Italy. <sup>2</sup>The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tokyo, Japan. ✉email: eugenio.lippiello@unicampania.it

$z$  of the two dominant variants. Furthermore the ratio between the two peak heights also provide information about the relative incidence of the two variants in virus spreading.

We perform this study using the incidence rate  $I(t)$  for SARS-CoV-2 in Italy where three Variants of Concern (VOC) have dominated in three different temporal windows, as highlighted in Table 1 using the information present in the Bulletin (No. 3 to 21) of Istituto superiore di Sanità ([www.epicentro.iss.it](http://www.epicentro.iss.it)) (see also Fig. 3report-n.21).

Several studies<sup>6,13–15</sup> have measured, in different geographic regions, a value  $\bar{z}$  of the Omicron variant significantly shorter than the value measured for previous variants Alpha and Delta. Because of this observation, many countries have applied a reduction of the duration of the quarantine period ([www.ecdc.europa.eu](http://www.ecdc.europa.eu)). This is in agreement with the evaluation of the intrinsic value of  $\bar{z}$  using nucleotide sequences of SARS-CoV-2 viruses sampled in Denmark<sup>16</sup>, leading to a value of  $\bar{z}$  for the Omicron variant about 0.5–0.6 times smaller than the one measured for the Delta variant. Conversely, a study of infections among household members in Reggio Emilia (Italy) lead to an intrinsic value of  $\bar{z}$  which is about 6 days, with no significant difference among the three variants, Alpha, Delta e Omicron<sup>10</sup>. Our data appear more consistent with the result of<sup>16</sup> since we find a value  $\bar{z} = 3.2 \pm 0.8$  days to the Omicron variant respect to the value  $\bar{z} = 6.5 \pm 1$  days associated to the Delta one.

### The method

In this section, we overview the method considered in this study, details can be found in<sup>5</sup>.

The starting point is the renewal equation<sup>11,12,17</sup> providing the expected value of daily infected people on the  $m$ -th day,  $E[I(m)]$ , in terms of the past daily incidence

$$E[I(m)] = \sum_{j=0}^{m-1} R_c(j)w(m-j)I(j) + \mu(m), \tag{1}$$

where  $R_c(m)$  is the case reproduction number, representing the total number of infections induced on average by an individual infected on the  $m$ -th day,  $w(j)$  is the distribution of generation times, representing the percentage of infections induced at a time distance  $j$  from the infection and, finally,  $\mu(m)$  is the daily number of imported cases during the  $m$ -th day, i.e. infectors coming from outside the considered region. We assume that  $w(j)$  is a Gamma distribution,  $w(j) = (\tau^{-a} / \Gamma(a))j^{a-1} \exp(-j/\tau)$ , which depends on two parameters,  $a \geq 1$  and  $\tau > 0$ , and where  $\Gamma(a)$  is the Gamma function. The Gamma distribution is fully characterized by its average value  $\bar{z}$  and by its standard deviation  $\sigma$ , which are both functions of  $a$  and  $\tau$ ,  $\bar{z} = a\tau$  and  $\sigma = \sqrt{a}\tau$ . In Supplementary Information (SI) we show that similar results are found for a Weibull or a log-normal distribution (Figs. Suppl.3, 4). In SI, we also show (Fig. Suppl. 5) that  $\sigma$  weakly affects the reproduction number  $R_c(m)$  but, as deeply discussed in<sup>5</sup>, it remains an important parameter in defining the appropriate length of quarantine.

An analytical expression for the log-likelihood  $LL$  of the time series  $\{I(m)\}_{m=1,\dots,N}$ , for assigned sequences  $\{R_c(m)\}_{m=1,\dots,N}$ ,  $\{\mu(m)\}_{m=1,\dots,N}$ , and for given values of  $\bar{z}$  and  $\sigma$  has been obtained<sup>5</sup> under the hypothesis that the number of individuals infected on the  $m$ -th day is Poisson distributed. For fixed  $\bar{z}$  and  $\sigma$ , the best series  $\{R_c(m)\}_{m=1,\dots,N}$  and  $\{\mu(m)\}_{m=1,\dots,N}$  which maximize  $LL$  are finally obtained by generalizing the Markov-chain-Monte-Carlo method introduced to find the optimal parameters in epidemic models for seismic occurrence<sup>18,19</sup>.

In the following, we define  $LL^{best}(\bar{z}, \sigma)$ , the value of  $LL$  in correspondence to the best series  $\{R_c(m)\}_{m=1,\dots,N}$  and  $\{\mu(m)\}_{m=1,\dots,N}$  and we explore its dependence on the parameters  $\bar{z}$  and  $\sigma$ . The identification of  $LL^{best}(\bar{z}, \sigma)$  allows us to obtain also an accurate estimate of  $\sigma$ , which represents a measure of the duration of the infectious period and which is difficult to be obtained by contact tracing<sup>2,3</sup>. The numerical code is available for open access at [github-algorithm](#). The pipeline can be found in Fig. Suppl. 1.

Starting date ending date	12/28/20 6/6/21 (%)	6/5/21 7/19/21 (%)	7/17/21 8/30/21 (%)	8/28/21 10/11/21 (%)	10/9/21 11/22/21 (%)	11/27/21 1/10/22 (%)	2/5/22 3/21/22 (%)	4/9/22 5/23/22 (%)	5/14/22 6/27/22 (%)
Alpha	74.9	35.9	2.3	0	0.3	1.2	0	0	0
Beta	0.8	0.3	0	0	0	0	0	0	0
Gamma	6.3	7.2	0.4	0	0	0.1	0	0	0
Eta	1.2	0.6	0.1	0	0	0	0	0	0
Kappa	0.3	1.3	1.2	1.5	0	0	0	0	0
Delta	–	46.0	88.1	90.8	91.4	66.7	0.7	0.1	0
Omicron-BA1	–	–	–	–	–	15.5	63.7	4.4	5.5
Omicron-BA2	–	–	–	–	–	–	27.3	82.0	56.0

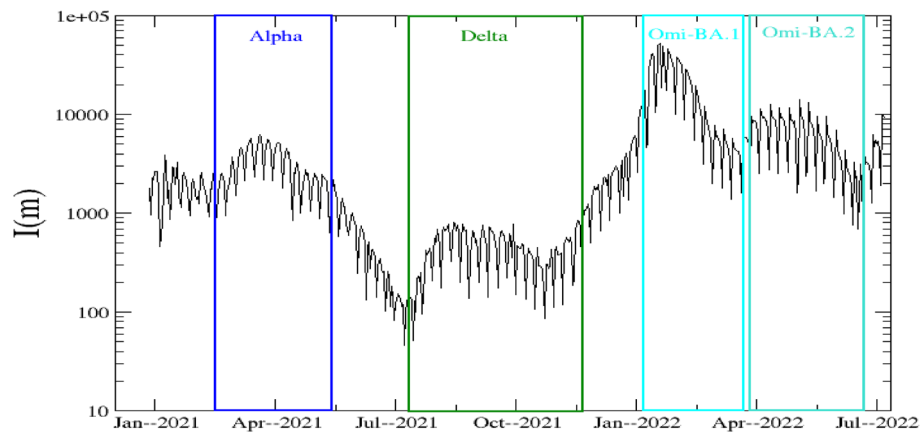
**Table 1.** Percentage of diffusion of major variants in Italy from 28 December, 2020, to June 27, 2022, based on weekly sampling and data provided by the I-Co-Gen platform software. Data are extracted from the reports titled “Prevalenza e distribuzione delle varianti di SARS-CoV-2 di interesse per la sanità pubblica in Italia”, from Report No. 3, dated June 25, 2021 up to Report No. 21, dated July 1, 2022, available at the official website of the Italian National Institute of Health (ISS) [www.epicentro.iss.it](http://www.epicentro.iss.it). Different rows correspond to different SARS-CoV-2 variants, different columns correspond to different temporal windows.

### Results

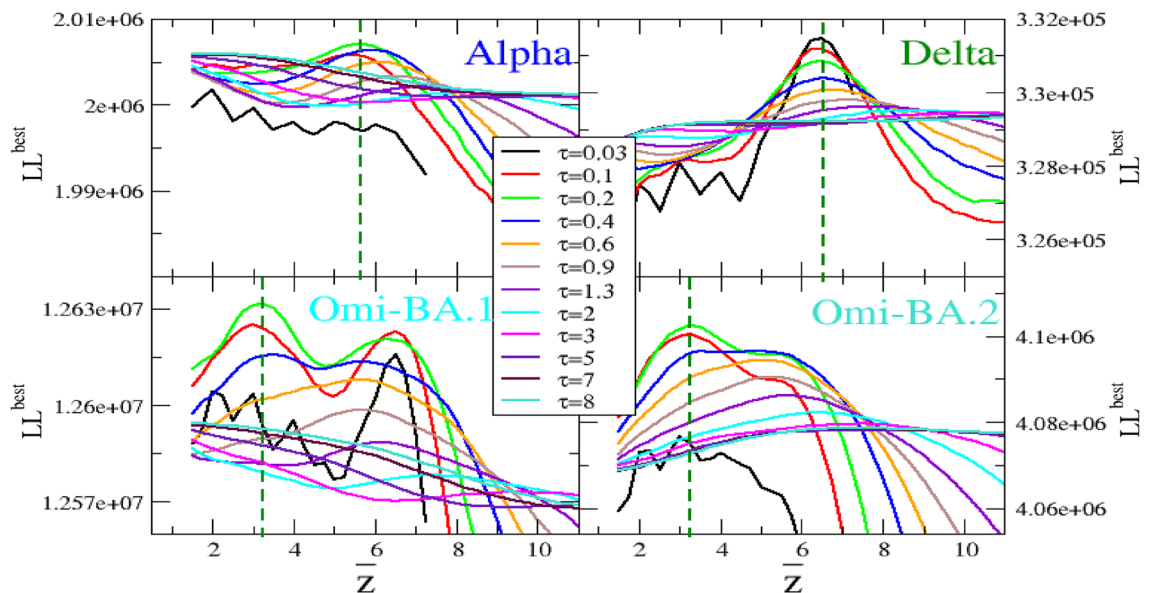
We consider data provided by the Department of Protezione Civile in Italy. More precisely we mainly consider data for the region Lombardy where the first outbreak of SARS-CoV-2 has been documented in Europe and which is characterized by a widespread diffusion of the disease since March 2020. In Fig. 1, we plot the daily incidence from January 2021. In the figure, we highlight the three main temporal windows which, according to the results of Table 1, are mostly characterized by the spreading of a specific variant. We have also identified two sub-windows where the spreading is mainly controlled by two different lineages of Omicron BA.1 and BA.2, respectively. It is evident that each temporal window corresponds to a different wave of Covid spreading with a distinct peak in  $I(m)$ .

We separately apply the procedure outlined in the previous section, restricting to data within each of the 4 temporal windows, which are classified as Alpha, Delta, Omicron-BA.1 and Omicron-BA.2. The evolution of the reproduction number  $R_c(m)$  and of the daily number of imported case  $\mu(m)$  is reported in Fig. Suppl. 2.

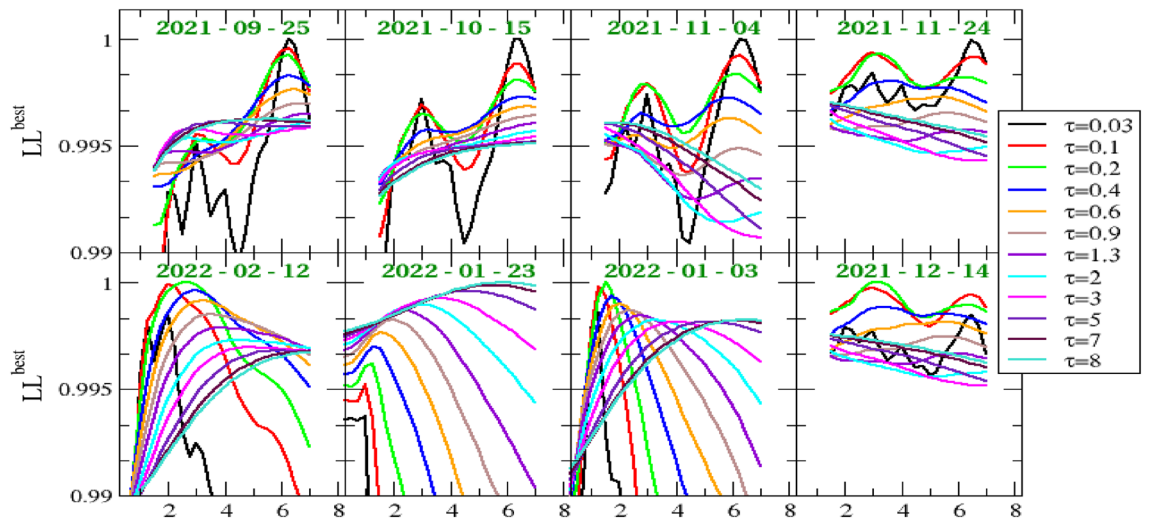
The values of  $LL^{best}(\bar{z}, \sigma)$ , for different choices of  $\bar{z}$  and  $\sigma$ , for each of the four temporal windows are plotted in a separate panel of Fig. 2. More precisely, we plot  $LL^{best}(\bar{z}, \sigma)$  versus  $\bar{z}$  and different values of the parameter  $\tau$  of the Gamma distribution. Results clearly show that during the Alpha window,  $LL^{best}(\bar{z}, \sigma)$  presents a clear



**Figure 1.** The daily incidence  $I(m)$  of SARS-CoV-2 in Lombardy from January 2021 up to June 2022. Colored boxes identify the four different temporal windows Alpha, Delta, Omicron-BA.1 and Omicron-BA.2 obtained from Table 1.



**Figure 2.** The log-likelihood  $LL^{best}(\bar{z}, \sigma)$ , evaluated for the temporal profile of  $R_c(m)$  which maximizes the likelihood for the daily incidence of SARS-CoV-2 in Lombardy, is plotted as a function of  $\bar{z} = a\tau$ . The four different panels correspond to the four temporal windows highlighted in Fig. 1. Different curves, in each panel, correspond to different values of  $\tau$ , which implies a different  $\sigma = a\sqrt{\tau}$ . The dashed green vertical line identifies the value of  $\bar{z}$  which provides the maximum value of the log-likelihood, in each panel.



**Figure 3.** The log-likelihood  $LL^{best}(\bar{z}, \sigma)$  is plotted as a function of  $\bar{z} = a\tau$ . Different panels correspond to different temporal windows of 60 days, with the initial day of each window reported on the top of each panel. The value of  $LL^{best}(\bar{z}, \sigma)$  has been divided by its maximum value  $LL^{best}(\bar{z}^{max}, \sigma^{max})$  in each panel, so that the maximum value of  $LL^{best}$  is normalized to 1 in each panel. Panels are organized in such a way that the initial time of temporal windows increases moving from left to right in upper panels and then continuous to increase moving in the lower panel from right to left. Different curves, in each panel, correspond to different values of  $\tau$ , which implies a different  $\sigma = a\sqrt{\tau}$ .

maximum for  $\bar{z} = 5.7$  days when  $\tau = 0.2$  days, leading to an estimate  $\bar{z} = 5.7$  days and  $\sigma = 1.1$  days, consistently with previous findings both in terms of serial interval and intrinsic generation time. During the Delta period the peak is even more pronounced at  $\bar{z} = 6.5$  days for  $\tau = 0.03$  days, consistently with previous results. Interestingly, during the Omicron-BA.1 window the maximum of  $LL^{best}$  at  $\bar{z} = 6.5$ , observed during the Delta window, is still present but is subleading, and the most relevant peak is present at a significantly smaller values  $\bar{z} = 3.2$  days for  $\tau = 0.2$  days. During the Omicron-BA.2 window the peak at  $\bar{z} = 3.2$  days is the only relevant one presented by  $LL^{best}$ . Results of Fig. 2 clearly show a significant reduction of the generation time of the Omicron variants, with an estimated value  $\bar{z} = 3.2$  days which is roughly half of the value estimated during the Delta period, in agreement with results of serial intervals and of Ref.<sup>16</sup>. On the other hand, we do not find significant differences for the value of  $\bar{z}$  between the two Omicron lineages BA.1 and BA.2. Figure 2 also gives  $\sigma = 0.8$  days for both Omicron lineages. This result, compared with  $\sigma = 1.1$  days measured during the Delta period, indicates that the standard deviation is similar for the different variants.

The analysis of Fig. 2 clearly indicates a significant reduction in the average generation time,  $\bar{z}$ , for the Omicron variants. Specifically,  $\bar{z}$  is roughly half the estimated value observed during the Alpha and Delta time windows. This finding is consistent with the results presented in Ref.<sup>16</sup>, where a similar conclusion was drawn based on the analysis of the serial interval distribution. Notably, we obtained similar estimates of  $\bar{z}$  and  $\sigma$  for the other Italian regions in each of the four temporal windows (Figs. Suppl. 7–10). Additionally, we demonstrate in SI that our estimate of  $\bar{z}$  is minimally affected by underestimates of the daily incidence rate  $I(m)$  due to unreported cases (Fig. Suppl. 6).

We remark that in the presence of two peaks of  $LL^{best}$ , if the range of parameters is not completely explored, it may occur that automatic procedures for log-likelihood maximization, based on Monte Carlo Markov Chains, could remain trapped in a relative maximum without reaching the global one. The result of Ref.<sup>10</sup> could be affected by this problem identifying as best model parameters the ones related to the Delta peak instead of those associated to the Omicron one.

In Fig. 3, we present the behavior of  $LL^{best}(\bar{z}, \sigma)$  as function of  $\bar{z}$  within temporal windows of a fixed duration of 60 days, with different starting days, ranging from the first one which is fully contained within the Delta window up to the last one which is fully inside the Omicron one. Results show that in the temporal window starting on 2021-09-25 (first upper panel) only the peak at  $\bar{z} \simeq 6.5$  is visible in  $LL^{best}$ . By shifting forward the starting time and considering a time window starting on 2021-10-15 (second upper panel) a subleading peak at  $\bar{z} \simeq 3$  appears. This second peak in  $LL^{best}$  therefore signals the presence of a new variant with a different  $\bar{z}$  in the first weeks of December 2021. This is fully consistent with the results of Table 1 indicating that the percentage of infections caused by the Omicron variant starts to be significant in the first weeks of December 2021. Moreover, consistently with Table 1, Fig. 3 shows that by further shifting forward the starting day (upper panels from left to right) the peak at  $\bar{z} \simeq 3$  becomes increasingly more relevant until it turns on the dominant one in the temporal window starting on 2021-12-14 (fourth lower panel). This is again consistent with the results of Table 1 indicating that the Omicron is the most relevant variant after the mid-December 2021. Keeping on shifting forward the starting time (lower panels from right to left) the peak at  $\bar{z} \simeq 3$  becomes more visible remaining the only relevant one in  $LL^{best}(\bar{z}, \sigma)$  in the time window starting on 2022-02-12. We remark that no clear indication can be extracted from  $LL^{best}(\bar{z}, \sigma)$  in the temporal window starting on 2022-01-23. In this case indeed a clear peak

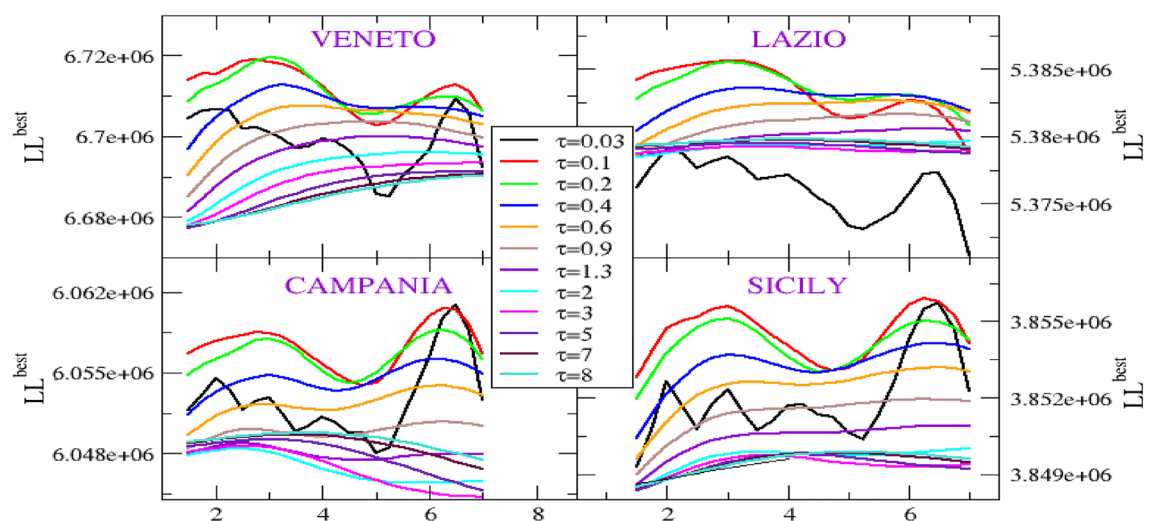
is not visible and the largest value of  $LL^{best}$  is obtained for the largest considered value of  $\tau = 8$  days, indicating that the standard deviation can be as large as 10 days and therefore does not allow us to obtain any information on  $\bar{z}$ . We have no clear justification for the very peculiar behavior of  $LL^{best}$  in this temporal window, which corresponds to the period when  $I(m)$  is in a fast decreasing phase. It could be possible that new infections within this time window are too few to extrapolate transmission parameters from  $I(m)$ .

In Fig. 4, we consider the behavior of  $LL^{best}(\bar{z}, \sigma)$  in different Italian regions during the Omicron-BA1 window. Results suggest the simultaneous presence of the two variants Delta and Omicron in all the considered regions. Indeed, in all regions the two peaks at  $\bar{z} \simeq 3$  and  $\bar{z} \simeq 6$  are clearly visible. However, the relevance of the two peaks is different among the different regions. Indeed, in some regions like Lazio, the Omicron variant clearly appears as the dominant one during the Omicron-BA.1 window. Conversely, in Campania the contagion appears still more controlled by the Delta variant whereas in Sicily the two variants appear to contribute in a similar way to SARS-CoV-2 diffusion. In Veneto, finally, one recovers a situation very similar to the one of Lombardy (Fig. 2) with a small predominance of the Omicron variant with respect to the Delta one.

## Conclusions

We have considered an epidemic model based on a renewal equation (Eq. 1) which depends on the transmission parameters  $R_c(m)$ , representing the time dependent case reproduction number, and on the parameters  $\bar{z}$  and  $\sigma$ , representing the mean value and the standard deviation, respectively, of the generation time distribution. We have used this model to describe the daily incidence rate of SARS-CoV-2  $I(m)$  in Italian regions during different temporal windows. More precisely, we have obtained the value of model parameters providing the best description of experimental data by using the log-likelihood maximization procedure introduced in<sup>5</sup>. In particular, we have separately considered data in four different temporal windows corresponding to periods when the diffusion of SARS-CoV-2 was mostly controlled by one of the four variants (Alpha, Delta, Omicron-BA1 and Omicron-BA2). We have found that  $\bar{z}$  during the Omicron windows was significantly smaller than, about one half of the value measured during Alpha and Delta windows, consistently with previous results about serial intervals<sup>6,13–15</sup> and an estimate of  $\bar{z}$  in Denmark<sup>16</sup>. By studying the behavior of the log-likelihood in different time windows, we find a clear indication of the presence of the Omicron variant in Italy since the first weeks of December 2021 with a diffusion becoming more and more relevant at later times. Our results are fully consistent with the relative diffusion of the different SARS-CoV-2 variants identified by sequencing provided by the I-Co-Gen platform software over the Italian territory. At the same time, we find that the standard deviation  $\sigma$  does not differ significantly in the different time windows.

Summarizing, our study shows that the adopted procedure can be very useful to identify, in about real time, changes in the transmission parameters of a virus that can be attributed to its mutations. We remark that this result can be obtained only from the daily number of infected individuals without any further information about the identification of the correct infector–infectee pair, ignoring the timing of symptom onsets as well as other details which are necessary to reconstruct the transmission chain in traditional approaches. More importantly, our approach does not need the support of laboratory analysis for genomic sequences, which is not always available. Accordingly, the procedure adopted in this manuscript could be particularly useful in the early stage of a new pandemic, or in the early stage of a new mutation, when the genetic information on the virus is not yet complete and genomic classification is not yet available. This procedure also allows one to monitor the evolution



**Figure 4.** The log-likelihood  $LL^{best}(\bar{z}, \sigma)$  is plotted as a function of  $\bar{z} = a\tau$  during the Omicron-BA.1 temporal window, for four Italian regions: Veneto (upper left panel), Lazio (upper right panel), Campania (lower left panel) and Sicily (lower right panel). Different curves, in each panel, correspond to different values of  $\tau$ , which implies a different  $\sigma = a\sqrt{\tau}$ .

of the standard deviation  $\sigma$ , which is an estimate of the duration of the infection period, an information complicated to be extracted by usual approaches based on genomic sequencing and contact tracing.

### Data availability

The datasets analysed during the current study are provided by Protezione Civile for the 21 Italian regions and collected in the repository data.

Received: 1 December 2022; Accepted: 6 July 2023

Published online: 17 July 2023

### References

- Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford Science Publications, 2002).
- Ganyani, T. *et al.* Estimating the generation interval for coronavirus disease (covid-19) based on symptom onset data, March 2020. *Eurosurveillance* **25**, 17. <https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257> (2020).
- Ferretti, L. *et al.* Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, 6491. <https://doi.org/10.1126/science.abb6936> (2020).
- Ferretti, L. *et al.* The timing of covid-19 transmission. *MedRxiv* <https://doi.org/10.1101/2020.09.04.20188516> (2020).
- Lippiello, E., Petrillo, G. & de Arcangelis, L. Estimating the generation interval from the incidence rate, the optimal quarantine duration and the efficiency of fast switching periodic protocols for covid-19. *Sci. Rep.* **12**, 4623. <https://doi.org/10.1038/s41598-022-08197-x> (2022).
- Backer, J. A. *et al.* Shorter serial intervals in sars-cov-2 cases with omicron ba.1 variant compared with delta variant, the netherlands, 13 to 26 december 2021. *Euro. Surveill.* **27**(6), 2200042. <https://doi.org/10.2807/1560-7917.ES.2022.27.6.2200042> (2022).
- Champredon, D. & Dushoff, J. Intrinsic and realized generation intervals in infectious-disease transmission. *Proc. Biol. Sci.* **282**, 1821. <https://doi.org/10.1098/rspb.2015.2026> (2015).
- Ali, S. T. *et al.* Serial interval of sars-cov-2 was shortened over time by nonpharmaceutical interventions. *Science* **369**(6507), 1106–1109. <https://doi.org/10.1126/science.abc9004> (2020).
- Park, S. W. *et al.* Forward-looking serial intervals correctly link epidemic growth to reproduction numbers. *Proc. Natl. Acad. Sci.* **118**, 2. <https://doi.org/10.1073/pnas.2011548118> (2021).
- Manica, M. *et al.* Intrinsic generation time of the sars-cov-2 omicron variant: An observational study of household transmission. *Lancet* <https://doi.org/10.2139/ssrn.4068368> (2022).
- Kermack, W. O., McKendrick, A. G. & Walker, G. T. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Char.* **115**(772), 700–721. <https://doi.org/10.1098/rspa.1927.0118> (1927).
- Grassly, N. C. & Fraser, C. Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477. <https://doi.org/10.1038/nrmicro1845> (2008).
- Brandal, L. T. *et al.* Outbreak caused by the sars-cov-2 omicron variant in Norway, November to December 2021. *Euro. Surveill.* **26**(50), 2101147. <https://doi.org/10.2807/1560-7917.ES.2021.26.50.2101147> (2021).
- Song, J. *et al.* Serial intervals and household transmission of sars-cov-2 omicron variant, South Korea, 2021. *Emerg. Infect. Dis.* **28**(3), 756–759. <https://doi.org/10.3201/eid2803.212607> (2022).
- an der Heiden, M., & Buchholz, U. Serial interval in households infected with sars-cov-2 variant b.1.1.529 (omicron) are even shorter compared to delta. *Epidem. Infect.* **To appear** (2022).
- Ito, K., Piantham, C. & Nishiura, H. Estimating relative generation times and relative reproduction numbers of omicron ba.1 and ba.2 with respect to delta in Denmark. *MedRxiv* <https://doi.org/10.1101/2022.03.02.22271767> (2022).
- Fraser, C. Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS One* **2**(8), 1–12. <https://doi.org/10.1371/journal.pone.0000758> (2007).
- Bottiglieri, M., Lippiello, E., Godano, C. & de Arcangelis, L. Comparison of branching models for seismicity and likelihood maximization through simulated annealing. *J. Geophys. Res. Solid Earth* **116**, B2. <https://doi.org/10.1029/2009JB007060.B02303> (2011).
- Lippiello, E., Giacco, F., de Arcangelis, L., Marzocchi, W. & Godano, C. Parameter estimation in the ETAS model: Approximations and novel methods. *Bull. Seismol. Soc. Am.* **104**(2), 985–994. <https://doi.org/10.1785/0120130148> (2014).

### Acknowledgements

E. L. and G. P. acknowledge support from project PRIN201798CZLJ. L. de A. acknowledges support from project PRIN2017WZFTZP. E.L. and L. de A. acknowledge support from VALERE project E – PASSION of the University of Campania “L. Vanvitelli”. G. P. acknowledges support from JSPS Kakenhi (B) 23H03358. All the authors acknowledge Jiancang Zhuang for the useful discussion.

### Author contributions

E.L. conceived the experiment and contributed to writing-reviewing and editing, data curation and validation, G.P., S.B. and L.A. contributed to writing-reviewing and editing, data curation and validation. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-38327-y>.

**Correspondence** and requests for materials should be addressed to E.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023